caBIG
Tissue Banks and Pathology Tools Workspace (TBPTWS)
Requirement Specifications Survey

I. Respondent Contact Information

Center:
Contact Name:
Contact e-mail:
Role (e.g. developer, adopter):

II. Document Purpose

The purpose of this document is to collect information regarding the specifications of existing specimen bank data management systems and the perceived requirements of any new system that would be developed and adopted for the cancer Biomedical Informatics Grid (caBIG). In order to minimize the time and effort required to collect pertinent information, a series of guided responses are provided which should be answered as indicated. In the event that the options provided do not adequately characterize features of the data management system, the respondent is asked to provide brief details regarding the unique aspects of their system. **All information obtained from this survey will be kept confidential and will only be distributed in de-identified or aggregate form**. This information will be utilized by the caBIG TBPTWS development team to guide the construction of a data management system that can be easily deployed or adopted by all caBIG members. Prior to the onset of building this system, a formal "Requirements Specification" technical document will be produced and will be available for review and comment.

III. Scope of Specimen Bank

A. Please indicate the nature of the specimen bank served by your data management system (circle all that apply):
    1. Limited specimen bank support for a single clinical trial
    2. Specimen bank support for multiple clinical trials, same organ system
    3. Specimen bank support for multiple clinical trials, multiple organ systems
    4. General archival specimen bank (banked specimens not tied to specific trials)
    5. Specimen registry (specimens tracked but not physically held)
    6. Specimen distribution (collection and distribution, but no banking)
    7. Other (please describe below):

B. Please indicate the approximate number:
    1. Number of independent protocols used for specimen collection

2. Total number of participants registered
3. Total number of specimens banked
4. Annual specimen accrual
5. Annual number of specimens distributed

C. Please indicate the type of specimens collected:
   1. Frozen Tissue Specimens
   2. Paraffin Blocks from Surgical Pathology Service (Physically Held)
   3. Paraffin Blocks form Surgical Pathology Service (Registry Only)
   4. Lavage Specimens
   5. Serum and/or Plasma
   6. Urine
   7. Peripheral blood cell pellet
   8. Bone marrow aspirates
   9. Extracted DNA
   10. Extracted RNA
   11. Protein Lysates
   12. Other (please describe below):

D. Where are specimens collected:
   1. From a single site within the institution
   2. From multiple sites within the institution
   3. From multiple institutions
   4. From multiple sites within multiple institutions
   5. Other (please describe below):

E. What are the specimen / participant relationships:
   1. Single specimen collected from a single participant at one time
   2. Multiple specimens collected from a single participant at one time
   3. Multiple specimens collected from a single participant at multiple times
   4. Multiple specimens collected from a single participant at multiple times in multiple studies
   5. Other (please describe below):

F. Where are specimens stored:
   1. In a single central location
   2. In multiple, physically distinct locations within the institution
   3. In multiple, physically distinct locations in different institutions
   4. Specimens are registered but not stored in bank

G. Bank to Institution Relationships:
   1. Does the bank collect tissue for only one medical/research institution
   2. Does the bank collect tissue for multiple medical/research institutions (more than one IRB, etc)

H. What associated clinical data is collected with each specimen?
   1. Donor Demographics
   2. Pathology Diagnosis and Findings

3. Laboratory Data (Tumor Markers, etc) on Donor
4. Therapy History of Donor
5. Outcomes (Recurrence, Progression)
6. Patient Clinical Trials Activity
7. Other

I. Are participants followed to update any of the clinical data below?
   1. Past or Future Pathology Reports
   2. Laboratory Data (Tumor Markers, etc)
   3. Therapy History of Donor
   4. Clinical Status (Quality of Life)
   5. Outcomes (Recurrence, Progression)
   6. Vital Status
   7. Most recent follow up date
   8. Patient Clinical Trials Activity
   9. Other

J. What is the immediate source of the clinical data collected?
   1. Pathology Reports
   2. Laboratory Reports
   3. Clinical Questionnaires
   4. Outcomes/Oncology Registries
   5. Medical Record
   6. Clinical Trials Management Systems
   7. Other

K. What Identifiers are stored with the specimen?
   1. Tissue Bank "Accession" Number (Coded Number)
   2. Surgical Pathology LIS Accession Number
   3. Surgical Pathology LIS Accession Number and Block Letter
   4. Social Security Number
   5. Clinical Trial Participant ID Code
   6. Hospital Patient ID
   7. Other System ID (Describe)

IV. Inter-Bank Relationships

A. Please indicate data relationships between your specimen bank and other specimen banks with which you are aware.
   1. This bank is a stand-alone operation and does not interact with any other banks
   2. This bank is stand-alone but could potentially interact with other relevant banks (e.g. similar organ site banks at other institutions or other organ site banks at the same institution)
   3. This bank interacts (but no electronic data transfer) with other banks (How many?)
   4. This bank interacts using electronic data transfer with other banks (How many?)
   5. Other (please describe below):

B. If there is electronic data transfer between other banks, describe the nature of the data exchanged.
    1. HIPAA De-identified Data
    2. Patient Identified Data
    3. Inventory Data
    4. Demographic Data
    5. Pathology Data
    6. Outcomes Data
    7. Other Data

C. If there are tissue samples exchanged between banks, describe the nature and circumstance of these transactions.

V. Current Database System and Tools

Please circle all statements that apply.

A. What is the current nature of your data system:
    1. We have no electronic data system (written log books only)
    2. Spreadsheet or other non-relational electronic system
    3. Stand alone relational database (e.g. Access, 4D, Filemaker Pro)
    4. Commercial product (Name:)
    5. Multi-tiered database server with dedicated client software
    6. Multi-tiered database web server
    7. Other (please describe below):

B. What modes of data entry do you currently utilize:
    1. Manual entry of data
    2. Bar Coding
    3. Text scanning and encoding technology
    4. Manually merging of electronic data files
    5. Direct database to database interconnectivity
    6. Other (please describe below):

C. What is the current disposition of your data system:
    1. Have no system
    2. Not satisfactory. Wish to replace it as soon as possible
    3. Adequate. Would replace it if something better was available
    4. Satisfactory. Might replace it only if a newer system was substantially better
    5. Established. Would not / could not consider replacing the system
    6. Other (please describe below):

D. How many Information Technology FTEs support the operation of your data system?

E. How is metadata handled in the tissue bank:

1. There are no written data definitions
2. Data definitions, Date Entry and Validation Rules are written and available on paper
3. Data definitions, Data Entry and Validation Rules are written and available on line
4. Data definitions, Data Entry and Validation rules are incorporated in the tissue bank software

VI. System Access

A. Please indicate methods in which users access your data system:
   1. Directly from a workstation that hosts the database
   2. Through dedicated client software and intranet communication
   3. Through web-based intranet communication (single institution)
   4. Through web-based internet communication (multiple institutions)
   5. Other (please describe):

B. Please indicate the types of users that access your system:
   1. Clinical coordinators / Honest Brokers entering HIPAA-identified participant (Donor) data
   2. Bank personnel entering specimen tracking data
   3. Supervisors which edit data and insert new projects
   4. Administrators with read only / report access
   5. Research investigators querying for specimens
   6. Other (please describe):

C. Do different users have levels of read permissions in your system?

D. Do different users have levels of write (i.e data entry) permissions in your system?

E. Does your system track user access to the system?
   1. Yes
   2. No

F. Does your system log transactions:
   1. Logs data reads
   2. Logs data writes
   3. Logs data changes/edits
   4. There is no transaction logging
   5. Other (Describe)

G. Please describe any other unique access features of your system below:

VII. IRB and Patient Confidentiality

A. Under how many different IRB (Human Studies) protocols are specimens collected? If possible, please attach copies of these protocols and corresponding consent from language (as they pertain to specimen banking) as **Appendix C**.

B. Does your IRB make provisions for banking specimens for future, unspecified research?

C. Does your IRB make provision for aggregation and/or long term clinical follow up of tissue donors (participants).

D. Are HIPAA-defined participant identifiers stored in your system?

E. Are specimens ever distributed with HIPAA-defined participant identifiers?

F. Are objects (i.e. participants or specimens) de-identified (coded) in your system? If so, explain the method of de-identification below:

G. Does your facility maintain an NCI-issued certificate of confidentiality?

H. Are research results stored in your system?

I. Please describe below the encryption / security measures utilized by your system to prevent access to participant identifiers:

J. How would you rate your working relationship with your IRB:
   1. **Poor.** Seldom communicate with the IRB; Many outstanding policy conflicts
   2. **Fair.** Seldom communicate with the IRB; No outstanding policy conflicts
   3. **Good.** Regular communication with the IRB; No policy conflicts
   4. **Excellent.** Proactively working with the IRB to shape policies

K. As much as possible, please briefly describe scenarios where the specimen bank has had policy conflicts with the IRB or where matters of patient confidentiality have been problematic.

L. Who is responsible for the appropriate research use of banked tissue?

VIII. Distribution, Sharing, Material Transfer, and Intellectual Property (IP)

A. Does the Bank work with Tissue Utilization Committees? (How many?)

B. Who actually selects and approves the distribution of tissue to an investigator?

C. How are specimens "prioritized" for distribution in the tissue bank?

D. How does your tissue bank measure investigator feedback?

E. How does the bank "market" itself and its tissue to investigators?

F. Do you distribute specimens to extramural investigators who are named investigators on prospective collection studies?

G. Do you distribute specimens to extramural investigators who are not part of the original collection protocol or who are requesting specimens from your general specimen bank archive?

H. Do you have a standardized Materials Transfer Agreement for any specimen that is distributed extramurally? If so, please attach a copy of this agreement as **Appendix D**.

I. Do you distribute specimens to commercial entities?

J. How would you rate your working relationship with your Technology Transfer Office:
   1. **Poor.** Prohibited from distributing materials extramurally; Many outstanding policy conflicts
   2. **Fair.** Policies for material/data transfer developed ad hoc on a case by case basis
   3. **Good.** Standardized agreements available
   4. **Excellent.** Proactive in working with Technology Office to streamline issues surrounding material transfer and IP specifically related to human specimens and associated data

K. As much as possible, please list key IP issues that have been raised at your institution with regard to sharing specimens and associated data with extramural institutions.

L. Does your institution have an official policy on the release of pre-publication and post-publication data? If so, please describe:

IX. Data System Objects

For the purposes of this survey, 'Objects' are defined as physical entities about which data is collected and stored, usually in discrete data tables. Please indicate which objects are represented in your data system (note that the actual names of these objects may differ from system to system). In addition, please include your system's data schema as **Appendix A**.

A. *Studies* (Projects): A collection of participants and corresponding specimens that are collected under a uniform protocol and informed consent process

B. *Participants* (Donors): An individual from whom specimens are collected

C. *Sites* (Collection Sites): An institution or collection area within an institution where specimens are collected

D. *Collectors*: Clinical staff that collect specimens.

E. *Admissions* (Tissue Collection Event): An event in time that results in one or more collected specimens from a participant

F. *Specimens*: Biological material that is collected from a participant

G. *Segments*: Aliquot or subdivision of a single collected specimen

H. *Samples*: Molecular material (e.g. DNA or RNA) that is isolated from a specimen or segment

I. *Arrays*: An ordered collection of specimens, segments, or samples grouped as a single unit

J. *Investigators* (Research Projects): A researcher to whom a specimen, segment, sample, or array is distributed for laboratory investigation

K. *Distributions*: An event in time that results in one or more collected specimens, segments, samples, or arrays to be distributed to an investigator under a defined IRB protocol for a specific research project

L. *Users*: An individual who has access to the data system

M. *Other*: Use the format above to list other objects represented in your data system:

X. System Data Elements

A. Please attach as **Appendix B**, a list of system data elements in the following format (This can be a dump of the table structures of a database):
Table NameData Element NameData TypeControlled Values? Description

B. Please list any sources of common data elements or unified coding schemes employed by your system.

C. Does your system store other specialized data types (e.g. digital images)? Please specify and describe how they are used.

XI. Use Cases

Below is a list of representative use cases that may be commonly employed by a specimen banking data system. Please see section IX for definitions of representative objects. For each scenario, please indicate: 1=This functionality is not needed in the system; 2=This functionality is currently not employed in the system, but would be desirable; 3=This functionality is absolutely essential for the system.

A. Data Entry
    1. Register a new study
    2. Register a new site
    3. Register a new collector
    4. Register a new investigator
    5. Register a new user
    6. Participant data
        a. Register a new participant to a study
        b. Register an existing participant to a new study
        c. Enter new clinical data on existing participant
    7. Admission data

   a. Register a new admission for a new participant
   b. Register a new admission for an existing participant
   c. Enter pathology data for admission
8. Specimen data
   a. Register a new specimen for a new admission
   b. Register a new specimen for an existing admission
   c. Create a new specimen from an existing specimen
   d. Enter research data for a specimen
9. Segment data
   a. Create a new segment from a new specimen
   b. Create a new segment from an existing specimen
10. Sample data
   a. Create a new sample from an existing specimen
   b. Enter research data for a sample
11. Array data
   a. Create a new array from existing specimens / segments
   b. Create a new array from existing samples
   c. Enter research data for a specimen / segment array
   d. Enter research data for a sample array
12. Distribution data
   a. Distribute existing specimen
   b. Distribute existing segment
   c. Distribute existing sample
   d. Distribute existing array
13. Please list other specific data entry tasks required / desired for your system below:

B. Data Update
1. Update study data
   a. Close study
2. Update site data
   a. Close site
3. Update collector data
4. Update investigator data
5. Update user data
6. Participant data
   a. Update participant demographics
   b. Update / Withdraw participant consent
   c. Update participant clinical data
7. Admission data
   a. Update pathology data for admission
8. Specimen data
   a. Update specimen location
   b. Update specimen status (available, accessed, processed, etc.)
   c. Update specimen QA (histological review, etc.)
   d. Register specimen distribution (linked to create specimen distribution)

9. Sample data
    a. Update sample location
    b. Update sample status
    c. Update sample QA
    d. Register sample distribution (linked to create sample distribution)
10. Please list other specific data update tasks required / desired for your system below:

C. Data Querying
1. Query for specimens / samples by study
2. Query for specimens / samples by collection site
3. Query for specimens / samples by participant
4. Query for specimens / samples by Clinical / Pathological criteria
5. Query for specimens / samples by specimen attribute
6. Query for specimens / samples by research data
7. Query for specimens / samples by investigator use
8. Please list other specific data query tasks required / desired for your system below:

D. Other
1. Please list other specific tasks required / desired for your system below:

XII. The caBIG Virtual Specimen Repository

One potential goal of the caBIG initiative is to create a virtual specimen repository where institutions could exchange specimen inventory data, actual biospecimens, and research data generated from such specimens.

A. Is your bank part of such a multi-institutional virtual tissue bank today?

B. Below, please indicate whether any of the following issues will impede the progress toward this goal at your institution (1=significantly prevent, 2=may prevent, 3=can be resolved, 4=will not impede):

1. IRB / Human Studies concerns about sharing specimen data (e.g. creating a web-accessible specimen catalog)

2. IRB / Human Studies concerns about sharing specimens with other investigators for research studies not initially presented in the collection protocol / consent form

3. IP concerns about sharing specimens with extramural institutions

4. IP concerns about sharing research data generated from shared specimens

5. Competing scientific interests for use of specimens

6. Limited Information Systems support to create the required interfaces for inter-institutional data systems communication

7. Perceived loss of control of specimens/data

8. Please list below other specific restrictions that may limit the ability to share biospecimens and biospecimen data at your institution:

Appendix A. Please attach your system's data schema

Appendix B. Please attach a list of your system's data elements

Appendix C. Please attach language utilized by IRB protocols and consent form documents associated with specimen collection and banking

Appendix D. Please attach any standardized Materials Transfer Agreement utilized by your bank

Appendix E. Please attach examples of any administrative or client reports generated by your bank

XIII. FREE TEXT SECTION

A. Please provide a diagram identifying the main stakeholders in the tissue bank (IRB, Sponsoring Projects, Research Projects, Tissue Donors etc.) and their relationships between each other and the tissue bank.

B. Please provide a free text description of how the following activities occur in the tissue bank:

1. How is a typical Specimen Accessioned?

2. How does an investigator request tissue from the bank and how does that request become a formal order and an actual distribution?

3. How does the bank Q/A its inventory?